programa final

Encontro Nacional de Produtores e Usuários de Informações Sociais, Econômicas e Territoriais

INFORMAÇÃO PARA UMA SOCIEDADE MAIS JUSTA

III Conferência Nacional de Geografia e Cartografia

IV Conferência Nacional de Estatística

Reunião de Instituições Produtoras

Fórum de Usuários

Seminário "Desafios para Repensar o Trabalho"

Simpósio de Inovações

Jornada de Cursos

Mostra de Tecnologias de Informação

27 a 31 de maio de 1996
Rio de Janeiro, RJ BRASIL

IBGE 60 anos 1936-1996

Uma das maneiras de olhar o ofício de produzir informações sociais, econômicas e territoriais é como arte de descrever o mundo. Estatísticas e mapas transportam os fenômenos da realidade para escalas apropriadas à perspectiva de nossa visão humana e nos permitem pensar e agir à distância, construindo avenidas de mão dupla que juntam o mundo e suas imagens. Maior o poder de síntese dessas representações, combinando, com precisão, elementos dispersos e heterogêneos do cotidiano, maior o nosso conhecimento e a nossa capacidade de compreender e transformar a realidade.

Visto como arte, o ofício de produzir essas informações reflete a cultura de um País e de sua época, como essa cultura vê o mundo e o torna visível, redefinindo o que vê e o que há para se ver.

No cenário de contínua inovação tecnológica e mudança de culturas da sociedade contemporânea, as novas tecnologias de informação - reunindo computadores, telecomunicações e redes de informação - aceleram aquele movimento de mobilização do mundo real. Aumenta a velocidade da acumulação de informação e são ampliados seus requisitos de atualização, formato - mais flexível, personalizado e interativo - e, principalmente, de acessibilidade. A plataforma digital vem se consolidando como o meio mais simples, barato e poderoso para tratar a informação, tornando possíveis novos produtos e serviços e conquistando novos usuários.

Acreditamos ser o ambiente de conversa e controvérsia e de troca entre as diferentes disciplinas, nas mesas redondas e sessões temáticas das Conferências Nacionais de Geografia, Cartografia e Estatística e do Simpósio de Inovações, aquele que melhor enseja o aprimoramento do consenso sobre os fenômenos a serem mensurados para retratar a sociedade, a economia e o território nacional e sobre as prioridades e formatos das informações necessárias para o fortalecimento da cidadania, a definição de políticas públicas e a gestão político - administrativa do País, e para criar uma sociedade mais justa.

Simon Schwartzman
Coordenador Geral do ENCONTRO

Promoção

Fundação Instituto Brasileiro de Geografia e Estatística
IBGE
Fundação Instituto Brasileiro de Geografia e Estatistica
IBGE
Associação Brasileira de Estudos Populacionais
ABEP

Co-Promoção

Associação Brasileira de Estatística
ABE
Associação Brasileira de Estudos do Trabalho
ABET
Associação Brasileira de Pós-graduação em Saúde Coletiva
ABRASCO
Associação Nacional de Centros de Pós-graduação em Economia
ANPEC
Associação Nacional de Pós-graduação e Pesquisa em Ciências Sociais
ANPOCS
Associação Nacional de Pós-graduação e Pesquisa em Geografia
ANPEGE
Associação Nacional de Pós-graduação e Pesquisa em Planejamento Urbano e Regional
ANPUR
Sociedade Brasileira de Cartografia
SBC

Apoio

Federação das Indústrias do Estado do Rio de Janeiro
FIRJAN
Academia Brasileira de Letras
ABL
Conselho Nacional de Pesquisas
CNPq
Financiadora de Estudos e Projetos
FINEP
Revista Ciência Hoje

Institutos Regionais Associados

Companhia do Desenvolvimento do Planalto Central
CODEPLAN (DF)
Empresa Metropolitana de Planejamento da Grande São Paulo S/A
EMPLASA (SP)
Empresa Municipal de Informática e Planejamento S/A
IPLANRIO (RJ)
Fundação Centro de Informações e Dados do Rio de Janeiro
CIDE (RJ)
Fundação de Economia e Estatística
FEE (RS)
Fundação de Planejamento Metropolitano e Regional
METROPLAN (RS)
Fundação Instituto de Planejamento do Ceará
IPLANCE (CE)
Fundação João Pinheiro
FJP (MG)
Fundação Joaquim Nabuco
FUNDAJ (PE)
Fundação Sistema Estadual de Análise de Dados
SEADE (SP)
Instituto Ambiental do Paraná
IAP (PR)
Instituto de Geociências Aplicadas
IGA (MG)
Instituto de Pesquisas Econômicas, Administrativas e Contábeis
IPEAD (MG)
Instituto do Desenvolvimento Econômico Social do Pará
IDESP (PA)
Instituto Geográfico e Cartográfico
IGC (SP)
Instituto de Apoio à Pesquisa e ao Desenvolvimento "Jones dos Santos Neves"
IJSN (ES)
Instituto Paranaense de Desenvolvimento Econômico e Social
IPARDES (PR)
Processamento de Dados do Município de Belo Horizonte S/A
PRODABEL (MG)
Superintendência de Estudos Econômicos e Sociais da Bahia
SEI (BA)

Coordenação Geral

Simon Schwartzman

Comissões de Programa

Confege

César Ajara (IBGE)
Denizar Blitzkow (USP)
Jorge Marques (UFRJ)
Lia Osório Machado (UFRJ)
Mauro Pereira de Mello (IBGE)
Speridião Faissol (UERJ)
Trento Natali Filho (IBGE)

Confest

José A. M. de Carvalho (UFMG)
José Márcio Camargo (PUC)
Lenildo Fernandes Silva (IBGE)
Teresa Cristina N. Araújo (IBGE)
Vilmar Faria (CEBRAP)
Wilton Bussab (FGV)

Comissão Organizadora

Secretaria Executiva - Luisa Maria La Croix
Secretaria Geral - Luciana Kanham
Confege, Confest e Simpósio de Inovações
Anna Lucia Barreto de Freitas, Evangelina X.G. de Oliveira, Jaime Franklin Vidal Araújo, Lilibeth Cardozo R.Ferreira e Maria Letícia Duarte Warner
Jornada de Cursos - Carmen Feijó
Finanças - Marise Maria Ferreira
Comunicação Social - Micheline Christophe e Carlos Vieira
Programação Visual - Aldo Victorio Filho e Luiz Gonzaga C. dos Santos
Infra-Estrutura - Maria Helena Neves Pereira de Souza
Atendimento aos Participantes - Cristina Lins
Apoio
Andrea de Carvalho F. Rodrigues, Carlos Alberto dos Santos, Delfim Teixeira, Evilmerodac D. da Silva, Gilberto Scheid, Héctor O. Pravaz, Ivan P. Jordão Junior, José Augusto dos Santos, Julio da Silva, Katia V. Cavalcanti, Lecy Delfim, Maria Helena de M. Castro, Regina T. Fonseca, Rita de Cassia Ataualpa Silva e Taisa Sawczuk

Registramos ainda a colaboração de técnicos das diferentes áreas do IBGE, com seu trabalho, críticas e sugestões para a consolidação do projeto do ENCONTRO.

Draft: April 11,1996

The Development and Use of Longitudinal Microdata: The U.S. Census Bureau's Center for Economic Studies' Experience With Confidential Survey Microdata

by

Robert H. McGuckin and Arnold P. Reznek[1]

Abstract

This paper discusses lessons learned from the experience of the Center for Economic Studies (CES) of the U.S. Census Bureau with the development and use of micro-level longitudinal data sets. While the research has improved our understanding of the workings of the economy, we focus here on the practical aspects of the experience. Perhaps the primary lesson that has been learned is that programs providing analytic users access to the microdata enable the U.S. Census Bureau and the researchers to become partners working together to improve the basic information that society uses to make decisions. An access program is a simple, cost effective way to take advantage of the coincidence of interests of statistical agencies and analytic researchers from both academic and policy areas.

---

[1] Center for Economic Studies, U.S. Census Bureau. The views expressed herein do not necessarily reflect those of the U.S. Census Bureau.

1. Introduction

In a typical year, the U.S. Census Bureau, like statistical agencies around the world, carries out numerous surveys of firms and establishments. These data are routinely processed, tabulated, and published as cross-section aggregates -- the minimum level of aggregation being determined primarily by considerations of confidentiality. The effort is directed toward tabular output with surveys designed to produce aggregate point-in-time cross-section estimates for use in the National Income and Product Accounts (NIPA), input-output tables, business cycle indicators, and productivity measures for various sectors.

Historically, the microdata that form the basis for these aggregate statistics were considered expendable once the tabulations had been made.[2] Moreover, even though construction of the first longitudinal establishment-level microdata database at the U.S. Census Bureau -- known today as the Longitudinal Research Database (LRD) -- dates from the early 1980s, only within the last 2 or 3 years has it been designated as one of the core assets at the U.S. Census Bureau. This recognition coincides with sharply increased demands for microdata from analytic subject matter specialists in both the academic and policy communities.

This paper discusses lessons learned from the experience of the Center for Economic Studies (CES) of the U.S. Census Bureau with the development and use of micro-level longitudinal data sets. While CES research has generated a large number of basic insights about the workings of the economy, we focus here on the practical aspects of the experience.[3] We place particular emphasis on access by analytic users to the confidential microdata, the partnerships that researchers form with the U.S. Census Bureau when they work at CES, and the advantages of the program to both groups.

Widespread support for microdata panels from the research and policy communities is a relatively new phenomenon in the area of "economic" statistics -- the data obtained from business establishments and firms. The predominant pattern, until recently, has been that academic economists worked primarily with theory, and business and empirical economists worked primarily with macroeconomic models and aggregate data. Aside from practical considerations -- such as the high costs of using microdata for research and policy analysis before the arrival of low cost computer hardware and software -- analysts and policy makers hoped that the losses introduced by use of aggregate data were small. While researchers recognized that aggregate

---

[2] Microdata from the U.S. Census of Manufactures exists for five years, four of them during the great depression, prior to 1963.

[3] For summaries of substantive work at CES see McGuckin (1995), McGuckin and Jensen (1996), and the last several CES Annual Reports.

statistics meant the sacrifice of information[4], they hoped these losses were an acceptable price to pay for the simplicity and tractability that aggregate data affords.

A key reason for the increased appreciation for microdata is the growing awareness that the information loss from use of aggregate data is significant. In study after study at CES, researchers are finding that most of the variation in behavior of plants is observed within traditional sectorial aggregations -- industries, geography, size classes, etc.[5] This means that, for example, the establishment -- not the firm or industry -- is often the most appropriate unit of analysis for understanding economic performance. And, in many cases, the loss of information from use of firm- and industry-level aggregates leads to seriously distorted and incorrect conclusions.[6]

The growing awareness that answers to certain types of questions require longitudinal data is also an important aspect of the increased demands for microdata. It is virtually impossible to sort out cause and effect relationships without panel data.[7] Moreover, evaluation of the effects of policy interventions is seriously compromised without observations of the performance of the affected entity before and after the imposition of the policy.[8] It is also critical that a sample of "control" observations be included in the empirical model that evaluates the effects of the policy. Microdata often makes this aspect of the empirical specification feasible.

---

[4] See Theil (1954) for example.

[5] Similar results are being obtained for other countries. See, for example, Davis, Haltiwanger, and Schuh (1996), especially Chapter --, where establishment based statistics on job creation and destruction are compared for various countries.

[6] See, for example, McGuckin and Nguyen (1995b) where it is demonstrated that aggregation to the firm level obscures the relationship of ownership change and productivity change. As another example, Baily, Bartelsman and Haltiwanger (1994) show that aggregate statistics for the manufacturing sector provide an incorrect picture of the relationship of downsizing in productivity growth. There are many other examples from recent CES studies.

[7] See Bernard and Jensen (1996) and McGuckin, Streitwieser, and Doms (1995) for discussion of this point in the context of empirical examinations of exporter performance and technology adoption, respectively.

[8] See Olley and Pakes (1992) for an example of the application of longitudinal microdata to an industry undergoing policy change. Also see Jarmin (1995, 1996) for studies involving evaluation of a specific Government program.

The U.S. Census Bureau, particularly in the area of business and economic statistics, has become more supportive of direct access to microdata and the development of longitudinal panels.[9] Some of this support reflects the increased demands of analytic users. Some of it is due to increased sensitivity to customer needs and requirements as budgets for traditional statistical data products are reduced.

An important, and often overlooked, factor has been recognition of the importance of analytical subject matter research to a world-class statistical program. Until recently, most managers at the U.S. Census Bureau saw subject matter research as a drain on resources, rather than an activity with potential benefits. The products generated by the CES research program have increased support among Census officials for preservation of and research with the microdata. Analytic use of the microdata has generated invaluable insights and improved the quality and usefulness of U.S. Census Bureau survey programs. In addition, microdata panels are leading to new data products that can be marketed in traditional ways, and they have become an asset in attracting survey business. In short, it is becoming more widely acknowledged that the microdata are, in themselves, a valuable statistical agency asset.

The attitudes of the U.S. Census Bureau have also been swayed by recent experiences with the data access program at CES. Having witnessed literally hundreds of researchers from government, academia, and research organizations use the data without incident over the past ten years has eased concerns about protecting the confidentiality of the data. Nevertheless, such concerns are important and remain a central feature of the rules and procedures that guide CES operations.[10]

The CES program offers researchers access to sampling frames and general data collection programs that provide detailed and generally representative data for very detailed geographic areas, products, and industries. Such access is tightly controlled: Projects are chosen to ensure fairness and scientific merit, and the confidentiality of data is maintained. However,

---

[9] The "demographic" side of the U.S. Census Bureau is also feeling increased demands for local area microdata for which public use databases are difficult to construct because of disclosure concerns.

[10] There are well understood legal (Title 13 U.S.C) and moral obligations to protect the confidentiality of the data that survey respondents provide. We believe that without confidentiality protection, the survey response rates would decrease to unacceptable levels (at least in the United States).

research results are widely disseminated and peer review of research findings is fostered in many ways.

The CES access program ensures that the data are used to their fullest extent. This is not just a matter of the researchers exploiting data already collected. While access means a bigger proportion of the data become available for analysis, access also means that researchers effectively become efficient agents for change and quality improvement at the U.S. Census Bureau. By working with the underlying data, the researchers suggest improvements to the basic data collection programs and design and measurement changes far more effectively than is possible when their contact with the data is in the form of aggregative statistics used in endeavors far removed from the producers of the data -- the statistical agency.

While the information collected by the Census provides a very detailed and complete profile of the structure of the economy, it does not by itself provide sufficient detail to understand many research and policy problems.[11] In fact, in a substantial portion of research projects at CES researchers bring new data sets and merge them with the Census microdata. This not only improves individual research projects, it also provides important feedback to the U.S. Census Bureau on the quality of its data and possibilities for more efficient survey designs. A related recent development is survey sponsorship by research organizations on the basis of researchers identifying important gaps in the information provided by the U.S. Census Bureau, sponsor design in collaboration with U.S. Census Bureau professionals and using U.S. Census Bureau sampling frames, with analysis of the survey results (perhaps combining with other data) undertaken in collaboration with the sponsor at U.S. Census Bureau facilities.[12]

The paper is organized as follows. We begin by describing why analytic subject matter research is crucial to statistical agencies and why statistical agencies are unlikely to be able to obtain the benefits from subject matter research in the absence of an access program. We also argue that the cost of providing access is low at a statistical agency, which already has collected the basic microdata for other purposes, and can cover many of the costs of the program with user fees. To provide a concrete example of a program that works, we next describe the model for access developed by CES. The CES model offers a practical program that can be adapted for use in a variety of circumstances. Next, we briefly summarize some main themes from the

---

[11] See Davis, Haltiwanger, and Schuh (1996) and Jensen and McGuckin (1996) for discussions of the relatively small percentage of variation in plant behavior explained by the variables typically collected by statistical agencies.

[12] While such merged data sets often create problems of bias and representativeness, these problems are no worse, and probably less severe at the U.S. Census Bureau.

substantive research at CES. This serves as a background for a discussion of how access and the resulting analytic research has benefitted the U.S. Census Bureau. These benefits derive from creation of new data products, improvements and extensions to existing data programs, and, most recently, new survey business for the U.S. Census Bureau.

In spite of these successes, all is not perfect, and we conclude with some unsolved problems and possible changes for the future.

2. Analytic Subject Matter Research and Access Are Important to Statistical Agencies.

A capability in analytical subject matter research is essential for statistical agencies for many reasons. First, it provides the agency with an understanding of the problems with its data: Without using the microdata it is difficult to understand what is wrong with them or how to improve them. Second, with an analytic capability it is possible for a statistical agency to provide the public, in an non partisan way, with an understanding of the economic forces that affect them. Third, analytic work provides an understanding of the needs of customers, including both the executive and legislative branches of the government. Fourth, new data products are most often generated as a part of analytic research and problem solving. Thus, analytic research provides "essential" research and development services to the agency. Finally, analytical research can lead to new surveys since analysis helps to identify areas where information is incomplete and new data are needed for resolution of an issue. In fact, the U.S. Census Bureau has attracted several recent surveys because of the possibility of analytic work with the survey microdata, coupled with the LRD and the program for data access.

A comparison of the experience of demographic programs that deal with household and individual data with that of economic programs relying on business data illustrates the value of access to microdata. No known theorem argues that the supply side of the labor market is less important than the demand side. Yet, for many years, labor market researchers have had microdata on individual workers (the supply side of the market) available through public use files. Moreover, funds have been available for longitudinal household surveys such as the Survey of Income and Program Participation (SIPP). There are other examples, such as the National Longitudinal Survey, conducted outside the U.S. Census Bureau.

A key difference between the economic and demographic areas has been the significant presence of analytical researchers with access to microdata in the demographic areas, and the absence of them, until recently, in economic areas. Access to demographic data is related directly to the existence of public use microdata files, which are released and widely used. Researchers can be accommodated on the demographic side because the large sample sizes involved make it possible to release masked microdata without fear of disclosure. In contrast, business data is

characterized by small numbers of observations and extremely skewed size distributions at appropriate levels of aggregation and thus cannot support useful public use databases. See McGuckin and Nguyen (1990) and Fuller (19__). This difference is a key reason why analytic use of business data must rely on direct access programs, at least in the United States.[13]

The use of microdata panels is not restricted to large area demographic data. On the business side, longitudinal microdata have been used extensively by finance economists: They have ready access to panels of daily security prices and income and balance sheet information on publicly traded companies. Finance specialists have used these panels to understand the operation of security markets and to develop new financial products (e.g., options, index mutual funds and other "derivatives") allowing greater flexibility for businesses seeking capital, as well as providing investors with new tools to manage their risks. It is probably not a coincidence that finance, like labor economics, has been one of the more dynamic areas of social science research in the past 20 years.

Regardless of how it is accomplished, access to the data is essential if the benefits of microdata panels are to be fully realized. It might be argued that agencies could simply develop internal programs of analytic subject matter research and thereby limit the possibilities for access and exposure to disclosure risk. It is always true that an increase in the number of people who have access to data increases the possibilities of disclosure. There are two strong arguments against this reasoning. First, the empirical evidence is against it. The CES program has offered analytic researchers access to economic microdata for some time without incident. Moreover, there are good reasons to expect these good results to continue.

Second, statistical agencies are not likely to develop analytic research programs on their own. Aside from difficulties in attracting the best academic talent on a permanent basis, internal research programs tend to be either co-opted by the political process or starved because of inadequate funding. By developing independent research associates, the possibility for scientifically sound research is maximized.

---

[13] There are public use data files of establishment data outside the United States. For example, Millward (1993) discusses the British Workplace Industrial Relations Surveys (WIRS) and the conditions under which they are made available. However, the United States has a different legal environment, and the WIRS data sets are based on small samples. For large samples, it is still unlikely that public use files will be released in the foreseeable future in the United States .

Beyond this, public use samples do not allow researchers to link the public use data to data from other sources except at often unacceptable levels of aggregation.

An important ingredient of such a program is independence and wide dissemination of research results. Peer review of these research results preserves independence and helps CES to recruit quality staff members and research associates. In the absence of freedom to publish findings independently of the statistical agency, researchers are not likely to become a part of the program. While there is always some risk that CES research output could be confused with the agency's output and that the agency's reputation may suffer as a result, in practice this has not been a problem. We believe this is an important distinction that has made the CES program relatively popular with analytic researchers.

3. The Center for Economic Studies – A Vehicle for Access

The mission of CES is centered on three closely related activities: the creation and maintenance of longitudinal microdata, the development and implementation of procedures and mechanisms to provide access to the confidential microdata, and analytic subject matter research with the microdata. Access is instrumental in maximizing the value of longitudinal microdata panels and the procedures used to accomplish it are a major issue for many statistical agencies both in the U.S. and abroad. In this section we focus on the ways CES provides data access and satisfies the requirements of analytic customers, while meeting the legal, moral, and political requirements that the U.S. Census Bureau operates under.

Models for Access – CES Research Model. Under United States law, microdata from the U.S. Census Bureau data programs are confidential and may only be used for statistical purposes at secure sites by U.S. Census Bureau employees or by individuals who have obtained Special Sworn Status (SSS) from the U.S. Census Bureau. The law provides specific penalties for violations.[14] The primary reason is concern for the privacy rights of companies that fill out survey forms, but the publication (intended or not) of microdata would also reduce co-operation with data collection programs and cause response rates to fall. so the agency has practical reasons to ensure confidentiality.

Researchers who use confidential U.S. Census Bureau microdata must obtain SSS status. They are subject to the same prohibitions, rules, and legal penalties as regular U.S. Census Bureau employees and contractors. In fact, researchers probably face greater penalties than employees because of the potential for loss of their professional credentials, as well as denial of future access to the microdata if they disclose confidential information. Once they are sworn in, researchers work at a Census secure site -- a site where appropriate measures to ensure physical

---

[14] The relevant law is Title 13, U.S.C., Section 214. Violations are punishable with a fine of not more than $5,000 and imprisonment of not more than five years, or both.

security are in place and where a Census staff member is in charge. The researchers work directly with the individual or firm specific microdata required for their project. For most practical purposes they function just like regular staff researchers: They take part in internal research seminars and offer advice and discuss their projects with operating Division staffs and statistical professionals. Aside from developing and enforcing basic operating procedures, CES employees protect data confidentiality by performing "disclosure analysis" on any results the researchers wish to publish or remove from U.S. Census Bureau grounds. This analysis ensures that U.S. Census Bureau policies are met with respect to release of data.

It is worth noting in this regard that the range of information that can be released without violations of confidentiality is much greater for analytic results (e.g. regression coefficients) than for tabular data. Indeed, CES procedures suggest researchers minimize tabular output because secondary disclosure is very difficult: The operating divisions release as much tabular data as possible, which limits the amount of additional tabulations that can be released. Use of regression coefficients which offer protection (the r-squares are never close to 1.00) and qualitative descriptions is possible because CES stands ready to verify the specific statistical results reported.

In addition to obtaining SSS and paying of a laboratory fee to cover their project costs, research associates' projects must benefit the data collection activities of the U.S. Census Bureau. Meeting this criteria involves two basic considerations. First, proposed research projects must use U.S. Census Bureau microdata and have scientific merit. Second, as part of their projects, the researchers must be willing to try to improve the micro databases at CES, produce new data products, or provide the U.S. Census Bureau with recommendations for improving its data programs -- improved survey concepts or questionnaires, better survey processing procedures, new data products, etc.

CES employees support the research projects by becoming closely involved with them. CES staff researchers, augmented by a small computer staff, also undertake an internal program of database development and economic research independently and jointly with the "research associates" (as SSS researchers are called at CES). CES staff researchers are an important part of the project choice process and become agents for change within the U.S. Census Bureau. With the knowledge gained through their research and interaction with the research associates, CES employees are an important part of the transmission of research results within the U.S. Census Bureau. This is how the U.S. Census Bureau learns how to improve its data programs. We give

several examples in Section 5. But, we also emphasize that in our experience research associates themselves also often have substantial contacts with operating division personnel.[15]

Research Data Center Program. Since space and resources are limited at CES headquarters and it is often very expensive for researchers to relocate for the time required for major projects , we have begun to establish secure facilities, called Research Data Centers (RDCs), away from CES headquarters. The pilot RDC is the Boston Research Data Center (BRDC), established in January 1994 under a grant from the National Science Foundation (NSF). This RDC has proved successful, and final negotiations are underway to establish a new RDC this spring or summer at the Carnegie Mellon University Heinz School of Public Policy and Management in Pittsburgh. In addition, we have had serious talks with researchers and potential sponsors in Los Angeles, San Francisco, New York, Chicago, and Atlanta about setting up other RDCs. With the exception of San Francisco, each of these areas has a U.S. Census Bureau Regional Office capable of accommodating an RDC. The Regional Directors have been very supportive of the RDC concept, following the outstanding lead of Arthur Dukakis in Boston, who has been instrumental in the success of the first RDC. An important feature of the RDC program is its flexibility and ability to accommodate local variations. For example, an RDC can focus on specific studies of the regional economy (sometimes comparing it to the nation or other regions) and local economic issues. Or an RDC could develop specialties in areas like health, crime, or environmental issues. In fact, the Boston RDC has developed a concentration in environmental economics research. An RDC prospectus is now available to parties interested in creating new RDCs.

4. Subject Matter Benefits From Data Access

A major theme that has emerged from the CES research program is that heterogeneity in the distribution of business units is pervasive along a wide variety of dimensions. Firms differ dramatically, even within the same geographic areas and similar regions, and within four-digit industries and five-digit product classes as defined by the Standard Industrial Classification (SIC). Heterogeneity is observed across time as well as in the cross section. Not only does the growth process differ across firms, it is characterized by large discrete movements rather than smooth or continuous changes even for those establishments and firms in continuous operation. During any time interval, observed changes are "lumpy" and uneven. Some business units open and some grow, while others shrink and die. These facts raise the issue of what is the appropriate level of aggregation for analytic research.

---

[15] The number of CES permanent staff ranges from 20-30. With over 50 projects underway during any year, CES staff can not undertake all such contacts with the other operating divisions.

Research Without Representative Firm Assumptions. The representative firm assumption helps to reduce the myriad of economic activity to manageable proportions --and it provides confidentiality protection -- but it assumes that the behavior of all agents is alike. This assumption can be quite restrictive when there is substantial heterogeneity in the distribution of firms within a sector. The research shows that the behavior of firms and establishments varies greatly within typical sectoral classifications (it is "idiosyncratic"). This is true no matter what variables are analyzed (for example, output, employment, investment, or productivity), no matter what sectors are used to classify the data (for example, industry, size or location) and no matter what topics are examined (for example, merger policy, job turnover, business cycle analysis, research and development, energy consumption, pollution emissions, or pollution abatement expenditures). In the face of idiosyncratically behaving agents, aggregation error is introduced. Thus, a primary strand of research at CES is the evaluation of aggregation error.

Examining individual changes is necessary if particular components within an aggregate move differently from each other. An important example of the pitfalls in relying on aggregate data alone involves the relationship between productivity growth and employment change in U.S. manufacturing. In the aggregate, the manufacturing sector has increased productivity and shed workers over the past 10-15 years. The conventional wisdom is that rising productivity in manufacturing is due to firm downsizing. But the evidence suggests that this conclusion is misleading. Cross-sectoral reallocations of jobs (from manufacturing to services, for example) are small relative to reallocations within manufacturing. And within manufacturing, almost half of the productivity growth among continuously operating plants is associated with growing plants.

A similar situation can occur at the level of the firm: Even when the firm is the ultimate decision maker, data on the behavior of components of the firm, such as the plants it owns, may be required to understand the firm's performance. In recent work McGuckin and Nguyen (1995) show that the productivity enhancing effects of ownership change on manufacturing plants are obscured when firm level data is used in place of a model based on plant level observations.[16] The reason for this result appears to be that large multi-unit firms -- the kind most often examined by researchers because their data are public -- have diverse activities. Ownership changes in such companies typically involve large changes in the composition of the activities undertaken by the firm, and the use of firm level data can obscure the effects of the ownership change on the acquired properties.

---

[16] The plant-level model was estimated with a significant firm fixed effect so that characteristics of the firm are important determinants of plant performance. Thus, the failure of the model specified on firm-level responses is due to aggregation.

Both of these examples illustrate the need for longitudinal microdata for certain types of problems. It is virtually impossible to sort out the effects of ownership changes on performance unless one observes the economic unit before and after the ownership change (mergers, divestitures, leveraged buyouts, etc.). It is also impossible to compare and contrast the role of upsizing and downsizing plants without using longitudinal data to classify plants into each category.

In the absence of longitudinal data, identifying cause and effect relationships is problematic. For example, if well managed firms are the ones that adopt productivity enhancing technology, it is not possible from cross-section estimates that show intensive use of technology associated with high productivity to deduce whether the primary relationship is one of good businesses adopting technology or technology making businesses good. The problem is that both technology and productivity are correlated with an unobservable factor(s) called good management. See McGuckin, Streitwieser, and Doms (1996).

Heterogeneity Does Not Guarantee Aggregation Bias. It is important to note that the mere fact that establishments behave idiosyncratically is not sufficient to invalidate the use of aggregate data. Under certain conditions, the use of aggregate variables will introduce only negligible bias in an estimated relationship. Unfortunately, a long line of research has demonstrated that these conditions are quite restrictive. Even if interest centers only on aggregate responses to alternative policies (such as the effect of changes in pollution regulations, defense expenditure reductions, or tariff increases), responses will not be captured by simple linear functions of an average or representative firm if the responses of individual firms to changes are very different. In such cases, industry responses will be a weighted average of individual responses, and the weights can change over time.

Moreover, the effect of heterogeneity is not simply a matter of differences in firms and plants that continuously operate in an industry. Entry and exit decisions also generate aggregate industry responses that are not simple linear functions of the representative firm. Thus, the process determining survival is important for determining the proper level of data aggregation in a study. Work at CES by Olley and Pakes (1992), specifically investigating these issues, demonstrated significant errors in using aggregate data to estimate productivity relationships in telecommunications, an industry with substantial entry and exit.[17]

---

[17] Pakes and colleagues are constructing similar models to investigate the effects of policy changes (like energy tax increases) in the automobile industry. See McGuckin (1995) for a discussion and references.

This point is an important one: Recent empirical work at CES provides overwhelming evidence that not only is heterogeneity observable in cross-section data, but that plant and firm level responses to economic change are heterogeneous. In fact, most of the observed variation in the data is *within* industries. Moreover, the vast majority of this variation is *not* associated with traditional observables such as location, industry, size, age, or capital. Rather, this variation is associated with unobserved factors specific to the firm or business unit, many of which appear to be permanent attributes of the unit. Thus, linking basic operating data of the type typically collected -- shipments, value added, materials, labor, and capital -- is not sufficient for understanding the dynamics that govern economic growth. See Jensen and McGuckin (1996) for a summary of CES work. This body of work, as well as work from other sources (Bertin, Bresnahan and Raff 1992; Bresnahan and Raff 1991; Bresnahan and Ramey 1992) shows striking heterogeneity in the levels and movements of productivity, employment, growth, output, product structure, investment, and ownership change among establishments in similar markets, industries, and cohorts.

5. How Analytic Research Has Helped the U.S. Census Bureau -- Specific Examples

Aside from the importance of the body of research results illustrated in the previous section, data access has provided direct benefits to the U.S. Census Bureau. At the U.S. Census Bureau, research access has been an important factor in the development of the LRD and new data products such as job destruction and creation series that are derived from it. The research program has also helped a launch a major effort to improve economic classification systems in North America, has provided suggestions for improvements and extensions to data programs, and has increased the U.S. Census Bureau's survey business. These benefits can be anticipated for any statistical agency -- perhaps any data gathering organization -- with a microdata access program. The benefits flow from the partnerships formed with the analytic research community.

Research Databases. From the very beginning, researchers with access to the microdata have played a central role in developing the unique databases at CES. Research needs have been important factors guiding the development of these databases. In carrying out their projects, researchers provide essential contributions to the development and improvement of the databases. We illustrate with three of our databases -- The LRD, the Research and Development (R&D) Database, and the Worker Establishment Characteristics Database (WECD).

Longitudinal Research Database (LRD). The oldest and still the most used database at CES, the LRD consists of annual cost and output data on manufacturing establishments (plants) from the Census of Manufactures (1963, 1967, 1972, 1977, 1982, 1987, and 1992) and from the

Annual Survey of Manufactures (since 1972), linked to form an unbalanced longitudinal panel[18]. Over one million plants have appeared in the LRD in at least one year. Thus, the LRD is one of the most ambitious and comprehensive datasets available for the study of manufacturing. It has given rise to a large and varied body of research and policy analysis, and has formed the basis for several new statistical products (described below), the most notable of which is a series of annual (and quarterly) measures of job creation and job destruction.

The construction of the LRD was a major achievement. It grew from work in the late 1970's by the U.S. Census Bureau that was carried out under the direction of Richard and Nancy Ruggles of Yale University and funded by the national Science Foundation (NSF) and the Small Business Administration (SBA). In fact, CES was created in 1982 to facilitate research projects with the longitudinal data -- then called the Longitudinal Establishment Database (LED) -- and to improve, maintain, and update the basic data.

In the early years, work on the longitudinal panel focused on plants in continuous operation. This database of annual observations was called the time-series file. Its structure made it easy to carry out many types of analysis. For example, productivity growth was studied since plants were observed over the entire time interval. However, it soon became apparent that although balanced panels are very useful for many research issues, a balanced panel strategy for development of a longitudinal panel was inappropriate. Exits due to plant closings continually reduced the number of plants available for study. In addition to such losses from natural attrition, rotating sample designs also made big inroads in the size of the panel. Aside from sampling issues, the analysis of births and deaths has direct policy and research interest. Moreover, it is not enough to simply look at surviving plants, even if, for example, interest centers on production function elasticities.[19]

For these reasons, efforts shifted to creation of an unbalanced panel, called the LRD.[20] At the present time, CES research associates and CES regular staff are working together in a project

---

[18] A balanced panel would include data for all establishments in all years. The LRD is unbalanced because establishments are born and die, and because the ASM, as a sample survey, does not cover all establishments.

[19] See, for example, Olley and Pakes (1992).

[20] The name LRD was chosen to distinguish the new unbalanced panel from the LED. It also served to forestall controversy and complaints from U.S. Census Bureau managers who were concerned when industry totals from the LRD did not exactly match previously published figures. Such differences arise from new edits and different imputation procedures adopted for particular projects. See McGuckin and Pascoe (1988) for a more complete description of the LRD and its history.

-- actually a series of projects -- that will ultimately lead to a Longitudinal Business Database (LBD) that will cover virtually the entire economy, not just the manufacturing sector. The additional basic data will come from the Standard Statistical Establishment List (SSEL) -- the U.S. Census Bureau's master list of domestic establishments[21] -- and the quinquennial Censuses of Wholesale Trade, Retail Trade, Service Industries, Mineral Industries, and selected Transportation Industries.

Research and Development (R&D) Database. This database includes annual data from 1972 through 1993 from the RD-1 survey, a U.S. National Science Foundation-sponsored survey on firms performing R&D in the United States. The database is well suited for studies of firms' investments in technology. Several research projects have developed and refined this database. For more detail, see Adams and Peck (1994).

Worker Establishment Characteristics Database. The WECD is the result of a match between data from the 1990 Decennial Census long form (1 in 6 sample) and the _____ (Troske 1995). The motivation for constructing this database is that theoretical models in labor economics stress the importance of employer-employee matching in determining labor market outcomes, but most empirical work relies on either worker surveys with little information about employers or establishment surveys with little information about workers. With almost 200,000 workers matched to over 16,000 manufacturing establishments where they work, the WECD is the largest worker-firm matched data set in the United States. It has been used in several research projects on measuring the tendency of larger plants to pay higher wages (Troske 1994), wage determination based on worker characteristics and productivity (Hellerstein, Neumark, and Troske 1994); and the effect of technology use on the wages and skill mix of workers in plants (Doms, Dunne, and Troske 1994).

While the WECD contains a substantial number of workers and plants, it relates to a single year, covers only the manufacturing sector, and is not a random sample of workers or plants -- it contains more data on senior male workers in large plants in urban areas. Current research focuses on creating larger, more representative, longitudinal versions.

New Data Products. ***Job Creation and Destruction.*** A new book by this title (Davis, Haltiwanger, Schuh 1996) will appear in April 1996. As part of their research, the authors created annual and quarterly time serie s for job creation and destruction from 1973-1988 for a large number of sectors of the economy. The basic data series as well as an extended series for 1988-93 are now part of the products produced by CES. WE plan to continue this series for the next few years, but we anticipate that these data series, as well as extended series covering

---

21 See U.S. Census Bureau (1979).

nonmanufacturing sectors, will soon become a regular U.S. Census Bureau data product. Aside from benefitting the U.S. Census Bureau, this work has spawned the development of similar products at other statistical agencies in both the U.S. (BLS) and around the world.

The work by Davis, Haltiwanger, and Schuh has broad implications for economic research and policy making in many areas. The issues dealt with are fundamental and cut across a wide variety of fields in both macro and micro economics. For example, the high rates of job destruction documented in virtually every sector of the economy argue strongly that workers need the flexibility to adapt to changes in the location and skill requirements of jobs. The authors also show that the job reallocation rate is countercyclical: the job destruction rate shows greater cyclical variability than the job creation rate in the United States manufacturing sector. This fact runs counter to many business cycle theories and work is now underway to extend these insights. One possibility is that this work will provide a foundation for new leading indicators to improve forecasters ability to predict economic turning points. Part of our optimism in this regard is based on past successes, but a good portion can be traced to the explicit use of a non-representative firm paradigm as part of the research strategy.

Diversification Indexes. An index of manufacturing product diversification (Gollop and Monahan (1988,1991)) has several desirable properties and uses detailed product -level data from the LRD to measure diversification at the plant and firm level for the five Census years from 1963 to 1982. The index, together with some more recent evidence (Streitwieser (1991)), shows that over the last 30 years, firms became more diversified, but plants became more specialized in the products they produce.

Advanced Technology Products Series. CES has helped the Foreign Trade Division at the U.S. Census Bureau to develop improved statistics showing the volume of trade in advanced technology products (ATP). See McGuckin, Abbott, Herrick, and Norfolk (1992). Since January 1989, the ATP series has been a part of the U.S. Census Bureau's monthly trade statistics.

Support for Improved Economic Classification. One important set of longstanding measurement issues concerns the problems the Standard Industrial Classification (SIC) system has in trying to classify economic activity. Researchers working with regular staff provided several studies in this area. For example, a number of microdata examinations examined the feasibility and design of alternative systems (Abbott and Andrews 1990, McGuckin 1992, Mattey 1993). A vital part of this work involved the product heterogeneity component of the Gollop and Monahan (1991) manufacturing product diversification index. This quantitative evidence supports qualitative judgments of experts in classification. These computer- and staff-intensive computations could not have been carried out without the LRD.

Collaboration With U.S. Census Bureau Survey Programs. Researchers develop working relationships with U.S. Census Bureau production division staff members, as well as workers in other statistical agencies. These working relationships lead to informal (and sometimes formal) consultation and advice that benefits all parties involved. Although it is difficult to document all of these effects (particularly the informal ones), the following are some major projects and surveys to which CES has contributed.

Characteristics of Business Owners (CBO) Survey. The 1982 and 1987 CBO Surveys, sponsored by the U.S. Small Business Administration (SBA) and the U.S. Minority Business Development Agency (MBDA), provide data on the demographic and economic characteristics of business owners and the economic performance of their firms. Since the CBO over samples firms owned by minorities and women, it is particularly useful for studying small businesses owned by these groups. For more detail, see Nucci (1992). CES researchers worked with the CBO survey staff on the 1987 and 1992 surveys. In particular, research with the 1982 CBO microdata -- supported in part by the SBA and the MBDA -- resulted in a variety of improvements to the 1987 survey questionnaire and sampling scheme. CES also helped to develop longitudinal linkages between the 1982 survey data and the 1987 survey universe. Extensions of these linkages to the 1992 universe have begun.

Research and Development (R&D) Survey. Research projects by James Adams and William Long, working with SuZanne Peck, resulted in the construction and documentation of a longitudinal research data set for the R&D survey (Adams and Peck 1994). This research has generated suggested improvements to the sampling scheme and questionnaire for this survey (Adams and Champion 1992). Research associates Long and Bronwyn Hall are continuing to evaluate the data by comparing R&D data (from Form RD-1) with data from a matched set of 800-1,000 companies that also file R&D data with the Securities and Exchange Commission (filed on form 10-K).

Annual Survey of Manufactures (ASM). CES has been especially interested in the ASM, which is one of the two main sources of the LRD data (along with the Census of Manufactures). CES researchers have supplemented statistical research on the ASM by pointing out certain problems of bias resulting from ASM procedures. Davis, Haltiwanger and Schuh (1991) gave a complete description of the differences between published and sample (LRD) statistics in the ASM. McGuckin and Peck (1993) analyzed the effects on published statistics of the ASM rules under which establishments' industry classifications can change. Research associates from the Federal Reserve Board (FED) have examined adjustment factors for the ASM undercount. This work is expected to lead to a joint U.S. Census Bureau/FED/Bureau of Economic Analysis project to develop new adjustment factors.

Pollution Abatement Costs and Expenditures (PACE) Survey. The PACE database covers the years 1979-93 and is fully integrated with the Longitudinal Research Database, with annual additions as data becomes available. A new report by Mary Streitwieser (1996) describes the survey design and how it has evolved over time, evaluates the suitability of the data for research, and suggests changes to the survey. Future data users, in turn, will use the report and add their own suggested changes.

Increasing the U.S. Census Bureau's Survey Business

The CES access program provides survey sponsors with a way to bring their analytic capabilities to the survey microdata, even after linking their survey data to the existing Census databases. The capability to provide survey sponsors with direct access to survey microdata linked to the LRD helped generate new survey business for the U.S. Census Bureau. In 1994, the U.S. Census Bureau conducted the *National Employers Survey (NES)* survey of over 3,000 business establishments in the United States (about half in manufacturing) and their training practices. The NES provides information on who pays for training, how training and its returns are evaluated by firms, how much is spent on training versus recruitment of already trained workers, and who provides training - outside vendors and educational institutions. Survey analysis has been carried out at both the BRDC and CES headquarters, and the survey data for the manufacturing sector have been linked with the LRD to allow investigation of the effects of training on establishments' success over time.

6. Outstanding Issues

The CES research program has demonstrated the mutual benefits that accrue when researchers and the U.S. Census Bureau work together to exploit the microdata collected in its regular programs. In spite of this impressive record, there are ways to increase these benefits.

Support for Longitudinal Panels Should be Increased. The success of many studies at CES using longitudinal data sets like the LRD has led the U.S. Census Bureau to declare longitudinal economic microdata to be one of its core strengths. This is evidence that the U.S. Census Bureau and its customers have come to recognize the value of these panels in producing information valuable for public and private policy decisions. However, survey designs at the U.S. Census Bureau have yet to reflect this. For example, many manufacturing surveys, such as the Annual Survey of Manufactures, select their samples with probability proportional to size. The five-year ASM panels before 1979 included all plants of a firm if one establishment was selected for the firm. To reduce respondent burden, the 1979 panel and following panels no longer automatically select these other plants. Thus, the ASM includes a smaller percentage of

small establishments than large establishments and often does not include all of the plants of multi-establishment firms. In addition, editing and imputation routines typically use a limited number of adjacent years of data for a plant. Although this design and processing system has worked well for producing aggregate data, it is not ideal for longitudinal microdata research.

Need to Supplement Basic Census Data. Out of necessity, U.S. Census Bureau surveys are designed within limited resources to fill a variety of needs while minimizing respondent burden, and as a result much information is left out. Filling in these gaps can be accomplished in two ways: First, the basic microdata sets that form the LRD can be merged with data from both within and outside the U.S. Census Bureau. Particularly for merging internal data sets, it would be helpful if certain surveys were designed as supplements to other, more fundamental surveys, in much the same way as CPS supplemental questions are handled. Thus, for examples, surveys of energy use, capacity, pollution abatement expenditures, plant occupation distributions, and technology use would be designed as subsamples of the ASM. Also, if surveys were viewed in this way ownership changes might be more consistently carried forward across data sets.[22]

The EQW/NES training survey, sponsored by the DOE, provides another promising avenue. This establishment survey was designed to maximize the possibilities of matches with the LRD, thus providing time-series data on the matched establishments (and their firms) as well as additional cross-section information. The U.S. Census Bureau and other government agencies are now investigating other possibilities.

Need to Develop Methods of Handling Longitudinal Data -- not part of Survey Designs. Existing panel datasets such as the LRD are constructed by concatenating survey and U.S. Census Bureau data collected to serve another purpose. Linking complete business censuses offers few statistical problems. When, as is often the case, the data have been collected as part of a probability sample designed to develop a point-in-time estimate of some aggregate like GNP or industry output there are many problems in interpreting the results.

The lack of true probability designs raises tensions between survey statisticians and analytical researchers. Most panel data at CES comes from non-random designs. This is also true for linked data coming from several cross-section surveys, whether there is a longitudinal aspect or not. Such data cause many problems of representativeness that Government survey statisticians are often uncomfortable with. The problem is that such non-probabilistic datasets make statistical inference complicated. Moreover, the methodologies used by economists and other social scientists to handle such data are explicitly model based and rely heavily on

---

[22] Ownership changes have been a difficult problem in developing a longitudinal panel of R&D firms and in merging LRD plant-level data with R&D survey firm-level data.

economic theory. As such they are often outside the realm of standard statistical practices used by survey statisticians, most of whom spend their professional time dealing with the problem of point-in-time probability surveys.

The answer is not simply to try and develop true panel designs. They are very costly (attrition is a real problem) and difficult to process. In the face of budget constraints, this solution is unlikely to be adopted. Nor is it necessary.

The crux of the matter involves the proper use of non-random samples. Work on methodologies for use of non-random samples is crucial. The virtual absence of truly random probability samples bearing on the most important research issues in economic policy, as well as other social science issues, is one reason the issue of the proper analysis of non-random data will not go away. In such cases, one must be content with careful documentation, good theory to guard against spuriousness, and replication of the analysis with different datasets. Replication across a wide range of non-representative samples enables more general conclusions to be drawn. In fact, this is standard practice in the social sciences (Smith 1983). But it is not just social scientists that use non-random data. So do statisticians. Non-response problems are technically equivalent to (though perhaps less severe than) those facing a researcher who has only a non-random dataset to work with. The required adjustments complicate matters, and in some cases, are controversial. But these adjustments are workable (Laaksonen 1992).

In developing panels, statistical agencies need to be aware of these problems. More importantly, they must make it a priority to bring together -- amicably and productively --analytic users of the data with survey statisticians. This is particularly important because the professional standards and goals of the two groups can be quite different. Of course, in reporting the results of analysis using data, economists and other social scientists must carefully discuss the limitations of the data and not overgeneralize the results. This is, of course, not always done by researchers and such practices raise concerns with survey statisticians whether or not the data are from a panel.

7. A Concluding Comment

The CES program has resulted in a great deal of important and innovative analytic research; new published data products, suggestions for improving U.S. Census Bureau data programs, and a large and growing list of micro databases. It is also a key input of a new data collection strategy -- collection of data for a sponsor, while providing the sponsor with access to the survey microdata, linked to other microdata bases. These benefits flow from a strategy that exploits the natural coincidence of interests between analytic subject matter researchers, the U.S. Census Bureau, and survey sponsors, including governmental units. Since the understandings of the economy and insights into its operation rest heavily on longitudinal microdta, survey value

added is dependent on access to the microdata for analytic research purposes. In turn, the access program creates partnerships between the data collectors and users that generate dividends well beyond the sum of their individual contributions.

While there are still many problems and issues to tackle, based on our past experience they will be solvable if analytic researchers and statistical officials continue to work together to encourage partnerships such as those formed at CES.

References

Abbott, T.A. and S. Andrews (1990), "The Classification of Manufacturing Industries: An Input-Based Clustering of Activity," *Proceedings, Census Bureau Annual Research Conference*, pp. 26-45, *CES Discussion Paper #90-7.*

Adams, J.D. and E.J. Champion (1992), "Restructuring Research and Development Statistics at Census: A Blueprint for Change," paper presented to the Census Advisory Committee of the American Economic Association, October, 1992.

Adams, J.D. and S. Peck (1994), "A Guide to R&D Data at the Center for Economic Studies, U.S. Bureau of the Census," *CES Discussion Paper #94-9.*

Baily, M.N., E. Bartelsman, and J. Haltiwanger (1994), "Downsizing and Productivity Growth: Myth or Reality?," *CES Discussion Paper #94-4.*

* Baldwin, J. (19__)

Bernard, A. and J.B. Jensen (1994), "Exporters, Skill Upgrading, and the Wage Gap," *CES Discussion Paper #94-13.*

Berry, S., J.A. Levinsohn, and A. Pakes (1992), "Automobile Prices in Market Equilibrium," *Working Paper, Yale University Department of Economics.*

Bertin, A.L, T.F. Bresnahan, and D.M.G. Raff (1992), "Localized Competition and the Aggregation of Plant-Level Increasing Returns: Blast Furnaces 1929-1935," mimeo, Stanford University Department of Economics.

* Boyd, G., J.B. Jensen, and R.H. McGuckin (1996)

Bresnahan, T.F. and D.M.G. Raff (1991), "Intra-Industry Heterogeneity and the Great Depression: The American Motor Vehicles Industry 1929-1935," *Journal of Economic History*, Volume 51, pp. 317-331.

Bresnahan, T.F. and V. Ramey (1992), "Output Fluctuations at the Plan Level," *Working Paper #4105, National Bureau of Economic Research.*

Caballero, R. (1992), "A Fallacy of Composition," *American Economic Review*, Volume 82, pp. 1279-1292.

_____ and E.M.R.A. Engel (1992a), "Beyond the Partial Adjustment Model," *American Economic Review, Papers and Proceedings*, Volume 82, pp. 360-364.

_____ (1992b), "Microeconomic Adjustment Hazards and Aggregate Dynamics," *Working Paper #4090, National Bureau of Economic Research.*

_____ and J. Haltiwanger (1995a) "Plant Level Adjustment and Aggregate Investment Dynamics," *Brookings Papers on Economic Activity, Number 2*, pp. 1-54.

_____ (1995b), "Plant Level Adjustments and Aggregate Investment," unpublished working paper.

Davis, S. and J. Haltiwanger (1992), "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, Volume 107, pp. 819-864, *CES Discussion Paper #90-4.*

_____ and S. Schuh (91-1), "Published Versus Sample Statistics from the ASM: Implications fro the LRD," *Proceedings of the American Statistical Association Business and Economics Section, CES Discussion Paper #91-1.*

_____ (1996), Job Creation and Destruction, Cambridge, MA and London, England: The MIT Press, forthcoming.

Doms, M.E., Dunne, T., and Troske, K. (1994), "Workers, Wages, and Technology," mimeo, Center for Economic Studies, U.S. Bureau of the Census.

Dunne, T. and M.J. Roberts (1993), "The Long-Run Demand for Labor: Estimates from Census Establishment Data," *CES Discussion Paper #93-13.*

Dunne, T. and J. Schmitz (1992), "Wages, Employer Size-Wage Premia and Employment Structure: Their Relationship to Advanced-Technology Usage at U.S. Manufacturing Establishments," *CES Discussion Paper #92-15.*

* Gilchrist, S. (19__)

* Gollop, F. and J.L. Monahan (1988)

_____ (1991), "A Generalized Index of Diversification: Trends in U.S. Manufacturing," *Review of Economics and Statistics*, Volume 73, pp. 318-330.

Haltiwanger, J. (1993), "Proposal for Census Fellows Program," mimeo, U.S. Bureau of the Census.

Hellerstein, J., D. Neumark, and K. Troske (1994), "Wages, Productivity, and Worker Characteristics," mimeo.

Jarmin, R. (1995), "Using Matched Client and Census Data to Evaluate the Performance of the Manufacturing Extension Partnership," *CES Discussion Paper #95-7.*

* Jensen, J.B. (1996)

_____ and R.H. McGuckin (1996), "Firm Performance and Evolution: Empirical Regularities in the Microdata," mimeo.

Laaksonen, S. (1992), "Adjustment Methods for Non-Response and Their Application to Finnish Income Data, *Statistical Journal of the United Nations*, pp. 124-137.

_____ (1993), "Creation of the European Enterprise Panel: Strategies," *Proceedings, International Conference on Establishment Surveys.*

Lichtenberg, F. (1992), Corporate Takeovers and Productivity, Cambridge, MA: The MIT Press.

_____ and Siegel, D. (1987), "Productivity and Changes in Ownership of Manufacturing Plants," *Brookings Papers on Economic Activity*, Volume 3, pp. 643-673.

* Long, W.F. and B. Hall (19__)

Long, W.F. and D.J. Ravenscraft (1992), "The Financial Performance of Whole Company LBOs," *CES Discussion Paper #93-16.*

_____ (1993), "LBOs, Debt, and R&D Intensity," in The Deal Decade, M. Blair, ed., Washington, D.C.: The Brookings Institution, *CES Discussion Paper # 93-3.*

Mattey, J. (1993), Evidence on IO Technology Assumptions from the Longitudinal Research Database," *CES Discussion Paper # 93-8.*

McGuckin, R.H. (1992), "Analytic Use of Economic Microdata: A Model for Researcher Access With Confidentiality Protection," *Proceedings, International Seminar on Statistical Confidentiality*, pp. 83-97, *CES Discussion Paper #92-8.*

_____ (1995), "Establishment Microdata for Economic Research and Policy Analysis: Looking Beyond the Aggregates," *Journal of Business and Economic Statistics*, Volume 13, Number 1, pp. 121-126.

_____, T.A. Abbott, P. Herrick, and L. Norfolk (1989), "Measuring the Trade Balance in Advanced Technology Products," *Proceedings, 1989 American Statistical Association*, CES Discussion Paper #89-1.

McGuckin, R.H. and Nguyen, S.V. (1993), "Post-Reform Industrial Productivity Performance of China: New Evidence from the 1985 Industrial Census Data," *Economic Inquiry*, Volume 31, pp. 323-341.

_____ (1995a), "On Productivity and Ownership Change: New Evidence from the LRD," *RAND Journal of Economics*, Volume 26, Number 2, pp. 257-276, *CES Discussion Paper #93-15.*

_____ (1996), "Exploring the Role of Acquisition in the Performance of Firms: Is the 'Firm' the Right Unit of Analysis," mimeo.

_____ and S.H. Andrews (1991), "The Relationships Among Acquiring and Acquired Firms' Product Lines," *Journal of Law and Economics*, Volume 34, pp. 477-502, *CES Discussion Paper #90-12.*

McGuckin, R.H., S.V. Nguyen, and A.P. Reznek (1994), "The Impact of Ownership Change on Employment, Wages, and Labor Productivity in U.S. Manufacturing, 1977-1987," *Proceedings, NBER Conference on Income and Wealth, 1995, CES Discussion Paper #95-8.*

McGuckin, R.H. and G. Pascoe (1988), "The Longitudinal Research Database: Status and Research Possibilities," *Survey of Current Business*, Volume 68, pp. 30-37, *CES Discussion Paper #88-2.*

McGuckin, R.H. and S. Peck (1993), "Manufacturing Establishments Reclassified Into New Industries: The Effect of Survey Design Rules," *Journal of Economic and Social Measurement*, Volume 1, Number 2, *CES Discussion Paper # 92-14.*

McGuckin, R.H., M.L. Streitwieser, and M.E. Doms (1995), "The Effect of Technology Use on Productivity Growth," mimeo, CES.

Millward, N. (1993), "How British Labor Economists Benefit from Establishment Surveys," unpublished working paper.

Nucci, A.R. (1992), "The Characteristics of Business Owners Database," *CES Discussion Paper #92-7.*

Olley, G.S. and A. Pakes (1992), "The Dynamics of Productivity in the Telecommunication Equipment Industry," *Working Paper 3977, National Bureau of Economic Research, CES Discussion Paper #92-2.*

* Phillips, G. (19__)

Smith, B.T. (1983), "On the Validity of Inference from Non-Random Samples," *Journal of the Royal Statistical Society*, 146, 4, pp. 394-403.

Streitwieser, M.L. (1991), "The Extent and Nature of Establishment Level Diversification in Sixteen U.S. Manufacturing Industries," *Journal of Law and Economics*, Part 2, pp. 503-534, *CES Discussion Paper #90-8.*

_____ (1996), "Evaluation and Use of the Pollution Abatement Costs and Expenditures Survey Micro Data ," *CES Discussion Paper #96-1.*

Theil, H. (1954), Linear Aggregation of Economic Relations, Amsterdam: North Holland Publishing Company.

Troske, K. (1995), "The Worker-Establishment Characteristics Database," *CES Discussion Paper #95-10.*

_____ (1994), "Evidence on the Employer Size-Wage Premium From Worker-Establishment Matched Data," *CES Discussion Paper #94-10.*

Zakrajsek, Egon (1995), "Retail Inventories, Internal Finance, and Aggregate Fluctuations: Evidence from Firm-Level Panel Data," *CES Discussion Paper #95-9.*